UNITED STATES DEPARTMENT OF THE INTERIOR
BUREAU OF LAND MANAGEMENT
DENVER SERVICE CENTER

THE LIGHTER SIDE OF STATISTICS

MARCH 1985
REVISED FEBRUARY 1986

BY
KRIS ESHELMAN
SHIRLEY HUDSON
BOB MITCHELL
MIKE PELLANT
KAY THOMAS

THE LIGHTER SIDE OF STATISTICS

Statistical analysis is one of those processes that many field offices are using, are curious about, would like to use, or are trying to avoid. In all seriousness, statistics remind many people of the old saying "liars figure and figures lie." Believe it or not, statistics do have a place in BLM.

Principles of statistics are often used in private industry. Most people, if they are going to invest a large amount of time and scarce dollars, will not take a chance on something if they have a greater than 50 percent chance (odds) of being wrong. If they do take a chance at those odds, they are considered crazy. If they succeed, they are brillant and have a lot of luck.

The taxpayers' inventory and monitoring program is very similar to investing in private industry except taxpayers never want a large amount of time and money invested in something if there is a 50 percent or greater chance of being <u>wrong</u>. Since bureaucrats are never called brilliant and are usually thought of as crazy, we cannot operate with a 50 percent or greater chance of being wrong. So we have statistics, and thus mathematicians, statisticians, and biometricians to baffle and confuse everyone, and to calculate our chances of success and accuracy.

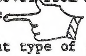
Statistics tell us the probability of success and let us know how confident we can feel about a value we have measured. The basic principles of statistics have made many casino operators rich men and have provided millions of dollars to Nevada's education system. (New Jersey does not count since it is not in the "real" West.) The "odds" are in favor of the casino at all times.

Now that we recognize the value of statistics, we can evaluate the reliability of the inventory and monitoring program, calculate how confident we are about the results of our work, and maybe avoid losing our hard-earned money next time we visit Reno, Las Vegas, or Panaca, Nevada.

The first objective of statistics is to place a range of values about a measured value (percent cover, production, etc.) and to state how confident we are (e.g., 80 percent or 90 percent) that the true value for that study sample is within that range. Statisticians call this calculating a confidence interval. The second objective is to pick a range of values around the mean, expressed as $\pm$ some percentage of the mean (statisticians say precision, and that $\pm$ is read plus or minus); pick a level of confidence; and then figure out how many samples we need to obtain that level of precision at that confidence. See, statistical language isn't completely confusing. The third objective is to evaluate the statistical significance of change that has occurred on a site or study over time.

Statistical terms that will be used are mean (average), precision, variance, standard deviation, coefficient of variation, and confidence interval. All you have to do is calculate them; I'll tell you how to use them. To make things easy, we are going to use a cookbook approach.

Now is a good time to get a pencil, a big eraser (or calculator with a $\sqrt{\ }$ key on it), a cup of coffee (to prevent sudden drowsiness) or another refreshment, and two aspirins if mathematics gives you headaches. Avoid alcoholic beverages! While you are up, grab an inventory or monitoring file and take out the data for one study. If you do not have a file handy, pull an example of density, frequency, production, or cover from Appendices 1, 2, 3, or 4. I will mark your spot while you are gone.



The first thing we have to determine is what type of study you have chosen to analyze. If it is cover, density, or production, you have to go to Part A; and if it is frequency, go to Part B. After you have mastered Parts A, B, and C (if you end up in C we are in trouble), you can use Part D to learn how to detect the statistical significance of any change.

Part A - Cover, Density, or Production

Please answer the following question: Did you (or somebody else) use more than one hoop, plot, or frame of uniform size to gather the cover, density, or production data; and are the data recorded plot by plot?

If your answer is no, then you are unable to continue further at least using Part A. Go to Section C for further instructions and suggestions. If your answer is yes, please continue.

Find Figure A_1 and take it out so you can use it. You should notice that there is a completed example on the right half of the page.

a. Pick a species (it is usually best to pick a key or dominant species) from your data and write the name in the appropriate blank. Fill in the attribute (i.e., density, production, etc.,) we are analyzing as well.

b. Count the number of hoops, plots, or frames that were sampled and enter the number in the space provided. Note the symbol in parenthesis (n) after the space; we will use this number later.

c. Enter the sample values (for the species you picked) in the vertical column (X) for each hoop, frame, etc. (plot 1, plot 2, etc.,).

d. Add up Column (X) and enter the answer in TOTAL (X)______.

e. Remember (n)(number of plots, frames, etc.)? Divide TOTAL (X) by (n). This number is the mean or average $(\bar{X})$ of your species for all plots. Fill in the $(\bar{X})$ blank and the $(\bar{X})$ column, all with the same value.

f. Subtract the $(\bar{X})$ column from the plot 1, plot 2, etc. values in the (X) column. Enter each answer in the two $(X - \bar{X})$ columns.

Figure A1

Species ________ State ________

Attribute ________ District ________

Date ________ Allotment ________

Study # ________

Number of plots, hoops, or frames ________ (n)

	(X)	-	$(\bar{X})$	=	$(X-\bar{X})$	x	$(X-\bar{X})$	=	$(X-\bar{X})^2$
Plot 1		-		=		x		=	
Plot 2		-		=		x		=	
Plot 3		-		=		x		=	
Plot 4		-		=		x		=	
Plot 5		-		=		x		=	
Plot 6		-		=		x		=	
Plot 7		-		=		x		=	
Plot 8		-		=		x		=	
Plot 9		-		=		x		=	
Plot 10		-		=		x		=	

TOTAL (X) ____ ÷ ____ (n) = ____ $(\bar{X})$ TOTAL A ____

____ (TOTAL A) ÷ ____ (n-1) = ____ (S^2)

$\sqrt{\_\_\_\_}$ (S^2) = ____ (S) ____ (S) ÷ ____ $(\bar{X})$ = ____ (CV)

INTERCEPT (n) and CV:

90% Confidence ____ % (∓ mean) 80% Confidence ____ % (∓ mean)

TO CALCULATE CONFIDENCE INTERVALS AROUND DATA

@ 90% CONFIDENCE

(100%- ____ (% ∓mean)) x ____ mean, density, cover, (or production value)= lower limit ____

(100%+ ____ (% ∓mean)) x ____ mean, density, cover, (or production value)= upper limit ____

@ 80% CONFIDENCE

(100%- ____ (% ∓mean)) x ____ mean, density, cover, (or production value)= lower limit ____

(100%+ ____ (% ∓mean)) x ____ mean, density, cover, (or production value)= upper limit ____

Species ABSP State YA

Attribute Density District 0470

Date 2/29/84 Allotment 1007

Study # 2

Number of plots, hoops, or frames 6 (n)

	(X)	-	$(\bar{X})$	=	$(X-\bar{X})$	x	$(X-\bar{X})$	=	$(X-\bar{X})^2$
Plot 1	20	-	16	=	4	x	4	=	16
Plot 2	14	-	16	=	-2	x	-2	=	4
Plot 3	15	-	16	=	-1	x	-1	=	1
Plot 4	12	-	16	=	-4	x	-4	=	16
Plot 5	18	-	16	=	2	x	2	=	4
Plot 6	17	-	16	=	1	x	1	=	1
Plot 7		-		=		x		=	
Plot 8		-		=		x		=	
Plot 9		-		=		x		=	
Plot 10		-		=		x		=	

TOTAL (X) 96 ÷ 6 (n) = 16 $(\bar{X})$ TOTAL A 42

42 (TOTAL A) ÷ 5 (n-1) = 8.4 (S^2)

$\sqrt{8.4}$ (S^2) = 2.9 (S) 2.9 (S) ÷ 16 $(\bar{X})$ = .18 (CV)

INTERCEPT (n) and CV:

90% Confidence 14 % (∓ mean) 80% Confidence 10 % (∓ mean)

TO CALCULATE CONFIDENCE INTERVALS AROUND DATA

@ 90% CONFIDENCE

(100%- 14 (% ∓mean)) x 16 mean, density, cover, (or production value)= lower limit 13.8

(100%+ 14 (% ∓mean)) x 16 mean, density, cover, (or production value)= upper limit 18.2

@ 80% CONFIDENCE

(100%- 10 (% ∓mean)) x 16 mean, density, cover, (or production value)= lower limit 14.4

(100%+ 10 (% ∓mean)) x 16 mean, density, cover, (or production value)= upper limit 17.6

g. Multiply one $(X - X)$ column times the other $(X - X)$ column. [We are actually squaring $(X - \overline{X})$]. Remember, a negative number times a negative number equals a positive number, so all of your answers will be positive. Enter each answer in the $(X - \overline{X})^2$ column.

h. Add up the $(X - \overline{X})^2$ column and enter the answer in TOTAL A.

i. Divide TOTAL A by $(n - 1)$ (your number of plots less 1) and enter in the (S^2) spaces. S^2 is the variance of your data.

j. Use your calculator to find the square root of (S^2) and enter it in the (S) blanks. (S) is the standard deviation and means very little to most people, but statisticians include it in most formulae they use.

k. Divide (S) by your mean $(\overline{X})$ and enter this innocent looking value in the (CV) space. You have just calculated the coefficient of variation (CV) for the species named "whatever."

You have now completed all the calculations necessary to determine the precision at a given level of confidence for your inventory or monitoring study.

Remove Figures A_2 and A_3. Note that A_2 is titled 90 percent confidence and A_3 is titled 80 percent confidence. Select one of the confidence figures. For most purposes in BLM, 80 percent confidence is adequate. You will notice on the figure(s) that the number of samples (plots) (n) is on the bottom, and the Coefficient of Variation (CV) is on the left. Do not worry about the numbers on the right yet. If your (n) figure is five or less, consider using Appendices 5 and 6 (enlarged versions of Figures A_2 and A_3).

Find your (n) number on the figure. Now find your (CV) on the left-hand side. Where do these points intersect? Staying between the curved lines, follow the curve all the way over to the right-hand side. What is the number? This is the precision or plus or minus percent of the mean figure. For example, if your precision was 15 percent and you used the figure for 80 percent confidence, it means that you can be 80 percent confident that your data for species "whatever" is within $\pm$ 15 percent of the actual mean.

Enter your precision value in the appropriate confidence (80 percent or 90 percent) level blank (under the title INTERCEPT (n) and CV). Now select the appropriate formula (same confidence percent) and calculate the upper and lower limits for the Confidence Interval (CI). Be sure to convert the percent values to their decimal equivalents, i.e., in Figure A_1 for the 90 percent confidence, the 100% - 14% becomes 0.86 whereas the 100% + 14% becomes 1.14. A confidence interval tells you that the true population value lies somewhere between the upper and lower limits 80% (at the 80% confidence level) or 90% (at the 90% confidence level) of the time.

In the example shown in A_1 besides the confidence interval for the mean we could calculate the density of Agsp to be 72,600/acre (16 x 43,560 ÷ 9.6) and enter this value into the formula to arrive at the confidence interval for the population value. At the 90% confidence level the range would be from 62,436/ac to 82,764/ac. Now use the other confidence level figure to calculate the CI.

That wasn't so hard! What happens if the boss sends you out to do a study and he expects you to be 80 percent confident that your key species data are ± 20 percent of the mean ($\bar{X}$). While in the field, you sampled nine plots and calculated a coefficient of variation (CV) of .50. Using Figure A_3, you find that your precision (±% of the mean) is ± 23 percent. (In the world of statistics, smaller precision values are "better" and statisticians [and I will, too] refer to those smaller values as a "higher level" of precision. Logical, isn't it?) Since your precision is lower than the boss wants, you must collect more data. How much more? Quite simple! Using Figure A_3, find the intersection for CV = .50 and (n) = 9. Now using the same (CV) value, move to the right until you reach the 20 percent "band." Keep going to the right until you hit a vertical line or "tick" mark. Now go down to the number of samples. This is the total number of plots (11) you have to sample to be 80 percent confident your data for species X will be within ± 20 percent of the actual study mean. Now using your data, subtract 5 percent from your precision (±% of the mean) and figure out how many plots to sample at this new level. REMEMBER, STATISTICS ARE MORE MEANINGFUL ON A SPECIES BY SPECIES BASIS!

I would recommend you try the section on frequency next. Do not be surprised if all the instructions are basically the same. There is little difference in how we look at frequency, density, cover, and production. In frequency you use transect data whereas in density cover, etc., you use plot data.

Detecting change is found in Part D.

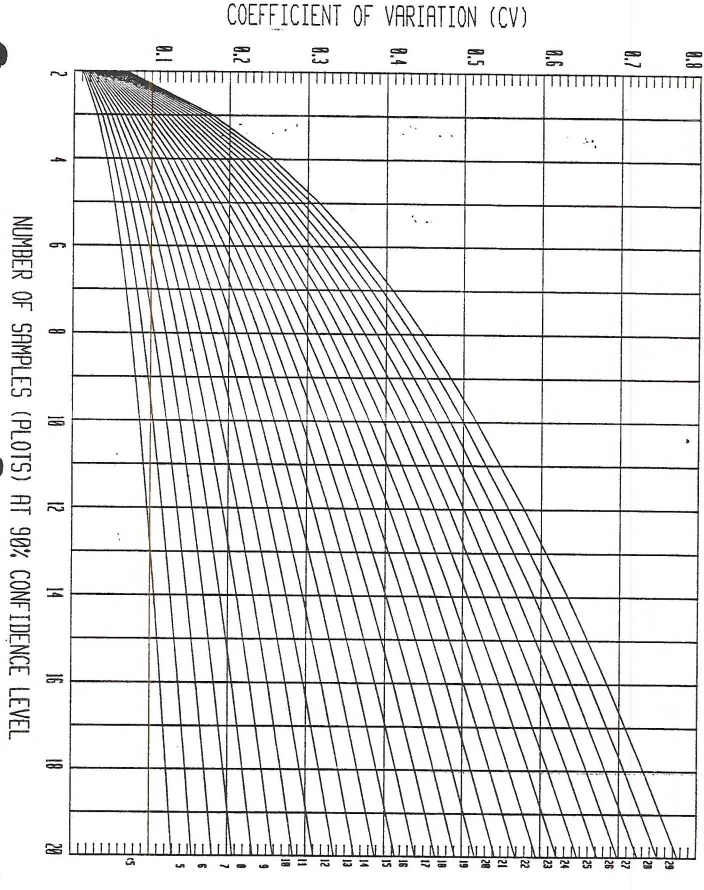
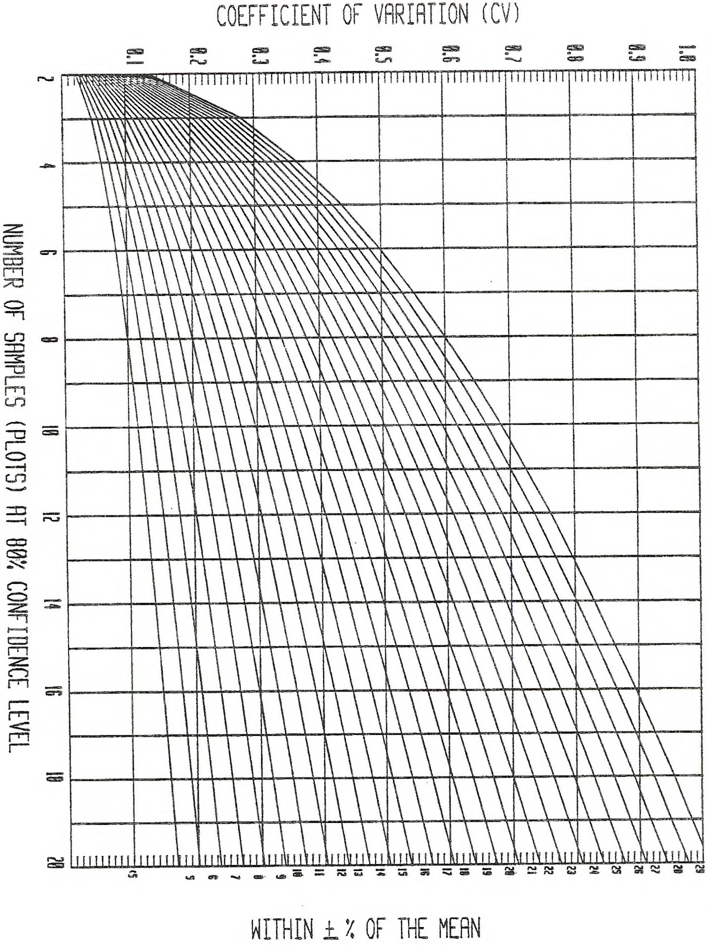

COEFFICIENT OF VARIATION (CV)
NUMBER OF SAMPLES (PLOTS) AT 90% CONFIDENCE LEVEL
WITHIN ± % OF THE MEAN

Figure A_2



COEFFICIENT OF VARIATION (CV)
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
NUMBER OF SAMPLES (PLOTS) AT 80% CONFIDENCE LEVEL
2
4
6
8
10
12
14
16
18
20
WITHIN ± % OF THE MEAN

Figure A_3

Part B - Frequency

I bet you are asking why frequency is different from density, cover, and production. Darn good question! Frequency is binomial data. That is, it is present or absent; either it is there or it is not; yes or no. Mathematically it is like having 0s (zeros) and 1s (ones); 1 for yes, 0 for no. There is a whole set of statistical formulas and theories about all these 0s and 1s. However, if we add up all the 0s and 1s for a transect, then we can use the same statistical formulas as for density, cover, and production. Furthermore, we get the same results that all the binomial theory and formula produce.

Instead of using plot by plot data we have to use transect totals. The first question we have to answer is "how many transects/subtransects (or belts) were run in your frequency study?". If your answer is greater than one, can you determine frequency values for each transect? If the answer is yes, we are okay. If your answer is no, you <u>may</u> have to go to Part C. In any case, please read on.

If you have only one transect, a statistician may not consider you his friend. Remember, they want people to do everything more than once. Don't worry though; if you have recorded your frequency data plot by plot, you have more options than most people. <u>If you did not record plot by plot, go to Part C.</u> If you sampled one transect plot by plot, you actually have a variety of subtransects. For example, 1 transect of 200 plots may be analyzed as 4 subtransects of 50 plots, 10 subtransects of 20 plots, 5 subtransects of 40 plots, etc. You can choose the configuration you want; and if it does not produce the results you want, try another configuration. <u>Usually</u> the more transects you have, the better your confidence and/or precision. Let's continue.

Find Figure B_1 and take it out so you can use it. You should notice that there is a completed example on the right half of the page.

a. Pick a species (it is usually best to pick a key or dominant species) from your data, and write the name in the appropriate blank. Fill in the attribute (frequency) we are analyzing as well.

b. Count the number of transects (or subtransects) that were sampled and enter the number in the space provided. Note the symbol in parenthesis (n) after the space; we will use this number later.

c. Enter the number of times (or plots) the species occurred (for the species you picked) in the vertical column (X) for each transect or subtransect (Transect 1, Transect 2, etc.).

d. Add up column (X) and enter the answer in TOTAL (X)____.

e. Remember (n) the number of transects (subtransects) etc.,? Divide TOTAL (X) by (n). This number is the mean or the average ($\bar{X}$) times your species occurred. Fill in the ($\bar{X}$) blank and the ($\bar{X}$) column, all with the same value.

f. Subtract ($\bar{X}$) from the Transect 1, Transect 2, etc., values in the (X) column. Enter each answer in the two ($X - \bar{X}$) columns.

g. Multiply one ($X - \bar{X}$) column times the other ($X - \bar{X}$) column. [We are actually squaring ($X - \bar{X}$)]. Remember, a negative number times a negative number equals a positive number, so all your answers will be positive. Enter the answer in the $(X - \bar{X})^2$ column.

h. Add up the $(X - \bar{X})^2$ column and enter the answer in TOTAL A.

i. Divide TOTAL A by (n-1) (your number of transects less 1) and enter in the (S^2) spaces. S^2 is the variance of your data.

j. Find the square root of (S^2) and enter it in the (S) blanks using your calcultor. (S) is the standard deviation and statisticians include it in most formulas they use.

k. Divide (S) by your mean ($\bar{X}$) and enter this innocent looking value in the (CV) space. You have just calculated the coefficient of variation for the species named "whatever."

You have now completed all the calculations necessary to determine the precision at a given level of confidence for your inventory or monitoring study.

Remove Figures B_2 and B_3. Note that B_2 is titled 90 percent confidence, and B_3 is titled 80 percent confidence. Select one of the confidence figures. For most purposes in BLM 80 percent confidence is adequate. You will notice on the figures that the number of transects/ subtransects (n) is on the bottom and coefficient of variation (CV) is on the left. Do not worry about the numbers on the right yet. If your (n) is five or less, consider using Appendices 7 and 8 (enlarged versions of Figures B_2 and B_3).

Find your (n) number on the figure. Now, find your (CV) on the left hand side. Where do these points intersect? Staying between the curved lines, follow the curve all the way over to the right-hand side. What is the number? This is the precision or plus or minus percent of the mean figure. For example, if your precision was 15 percent and you used the figure for 80 percent confidence, it means that you can be 80 percent confident that your data for species "whatever" is within $\pm$ 15 percent of the actual mean.

Enter your precision value in the appropriate confidence (80 percent or ninety percent) level blank (under the title INTERCEPT (n) and CV). Now using the correct formula (same confidence percent), calculate the upper and lower

Species ______ State ______
Attribute ______ District ______
Date ______ Allotment ______
Study # ______

Number of transects or subtransects ______ (n)

	(X)	-	$(\bar{X})$	=	$(X-\bar{X})$	x	$(X-\bar{X})$	=	$(X-\bar{X})^2$
Transect 1		-		=		x		=	
Transect 2		-		=		x		=	
Transect 3		-		=		x		=	
Transect 4		-		=		x		=	
Transect 5		-		=		x		=	
Transect 6		-		=		x		=	
Transect 7		-		=		x		=	
Transect 8		-		=		x		=	
Transect 9		-		=		x		=	
Transect 10		-		=		x		=	

TOTAL (X) ______ ÷ ______ (n) = ______ $(\bar{X})$ TOTAL A ______

______ (TOTAL A) ÷ ______ (n-1) = ______ (S^2)

$\sqrt{\_\_\_\_ (S^2)}$ = ______ (S) ______ (S) ÷ ______ $(\bar{X})$ = ______ (CV)

INTERCEPT (n) and CV:

90% Confidence ______ % (∓ mean) 80% Confidence ______ % (∓ mean)

TO CALCULATE CONFIDENCE INTERVALS AROUND DATA

@ 90% CONFIDENCE

(100%- ______ (% ∓mean)) x ______ (mean or frequency value)= lower limit ______

(100%+ ______ (% ∓mean)) x ______ (mean or frequency value)= upper limit ______

@ 80% CONFIDENCE

(100%- ______ (% ∓mean)) x ______ (mean or frequency value)= lower limit ______

(100%+ ______ (% ∓mean)) x ______ (mean or frequency value)= upper limit ______

Species HIJa State YA
Attribute Frequency District 0470
Date 10/87/48 Allotment 9999
Study # A3

Number of transects or subtransects 10 (n) 20 plots per subtran.

	(X)	-	$(\bar{X})$	=	$(X-\bar{X})$	x	$(X-\bar{X})$	=	$(X-\bar{X})^2$
Transect 1	16	-	14	=	2	x	2	=	4
Transect 2	13	-	14	=	-1	x	-1	=	1
Transect 3	15	-	14	=	1	x	1	=	1
Transect 4	12	-	14	=	-2	x	-2	=	4
Transect 5	8	-	14	=	-6	x	-6	=	36
Transect 6	15	-	14	=	1	x	1	=	1
Transect 7	16	-	14	=	2	x	2	=	4
Transect 8	20	-	14	=	6	x	6	=	36
Transect 9	10	-	14	=	-4	x	-4	=	16
Transect 10	15	-	14	=	1	x	1	=	1

TOTAL (X) 140 ÷ 10 (n) = 14 $(\bar{X})$ TOTAL A 104

104 (TOTAL A) ÷ 9 (n-1) = 11.6 (S^2)

$\sqrt{11.6 (S^2)}$ = 3.4 (S) 3.4 (S) ÷ 14 $(\bar{X})$ = .24 (CV)

INTERCEPT (n) and CV:

90% Confidence 13 % (∓ mean) 80% Confidence 10 % (∓ mean)

TO CALCULATE CONFIDENCE INTERVALS AROUND DATA

@ 90% CONFIDENCE

(100%- 13 (% ∓mean)) x 70 (mean or frequency value)= lower limit 61

(100%+ 13 (% ∓mean)) x 70 (mean or frequency value)= upper limit 79

@ 80% CONFIDENCE

(100%- 10 (% ∓mean)) x 70 (mean or frequency value)= lower limit 63

(100%+ 10 (% ∓mean)) x 70 (mean or frequency value)= upper limit 77

Figure B_1

limits for the confidence interval (CI). Be sure to convert the percent values in the formula to their decimal equivalents. For example, in Figure B_1, for 90% confidence, 100% - 13% becomes 0.87 whereas 100% + 13% becomes 1.13. The example in figure B_1 derived a frequency value for Hija by dividing the number of plots in which Hija occurred by the total plots sampled (140 ÷ 200 = .70 or 70%). A confidence interval tells you that the true population value lies somewhere between the upper and lower limits 80% (at the 80% confidence level) or 90% (at the 90% confidence level) of the time. Now use the other confidence level figure to calculate the CI.

That wasn't so hard! What happens if the boss sends you out to do a study and he expects you to be 80 percent confident that your key species data are ± 20 percent of the mean ($\overline{X}$)? While in the field you sampled 9 transects and calculated a coefficient of variation (CV) of .50. Using Figure B_3 you find that your precision (±% of the mean) is ± 23 percent. (In the world of statistics, smaller precision values are "better," and statisticians (and I will, too) refer to those smaller values as a "higher level" of precision. Logical, isn't it?) Since your precision is lower than the boss wants, you must collect more data. How much more? Quite simple! Using Figure B_3, find the intersection for CV = .50 and (n) = 9. Now <u>using the same (CV) value</u>, move to the right until you reach the 20 percent "band". Keep going to the right until you hit a vertical line or tick mark. Now go down to the number of transects/subtransects. This is the total number of subtransects (11) you have to sample to be 80 percent confident your data for species X will be within ± 20 percent of the actual study mean. Now using <u>your</u> data, subtract 5 percent from <u>your</u> precision (±% of the mean) and figure out how many transects to sample at this new level. REMEMBER, STATISTICS ARE ONLY GOOD ON A SPECIES BY SPECIES BASIS!

You have done so well you are hereby declared a member of the Royal Order of Befuddlers. If you haven't done so, I would recommend you try the section on cover, density, or production (Part A) next. Do not be surprised if all the instructions are basically the same. There is little difference in how we look at frequency, density, cover, and production. In frequency you use transect data whereas in cover, etc., you use plot data.

Detecting change is found in Part D.

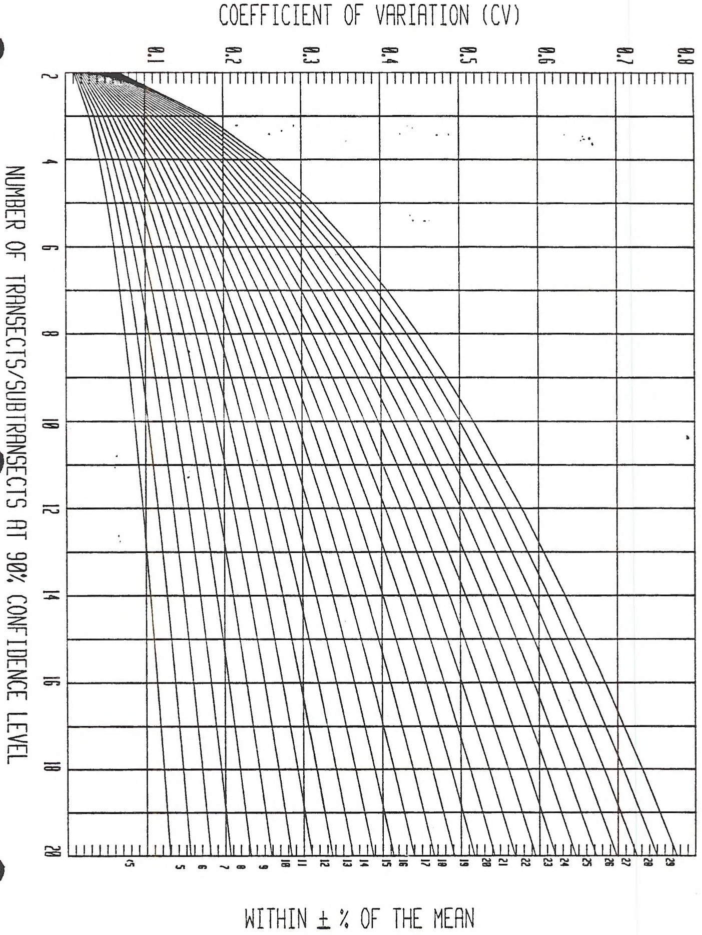

COEFFICIENT OF VARIATION (CV)
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
NUMBER OF TRANSECTS/SUBTRANSECTS AT 90% CONFIDENCE LEVEL
2
4
6
8
10
12
14
16
18
20
<5
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
WITHIN ± % OF THE MEAN

Figure B2

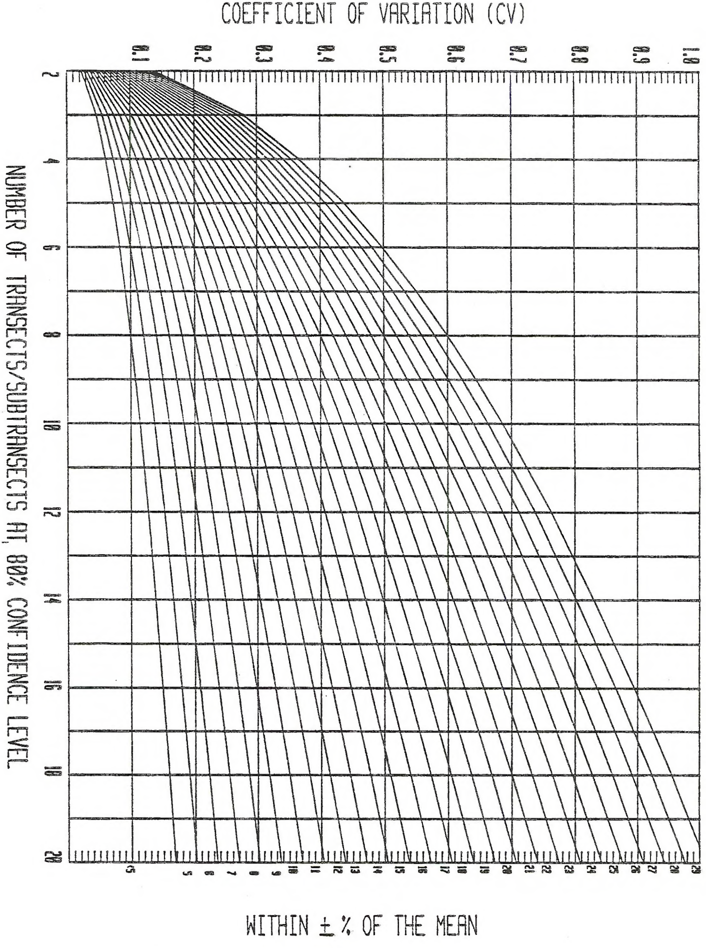

COEFFICIENT OF VARIATION (CV)
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
NUMBER OF TRANSECTS/SUBTRANSECTS AT 80% CONFIDENCE LEVEL
2
4
6
8
10
12
14
16
18
20
WITHIN ± % OF THE MEAN

Figure B_3

Part C - Never-Ever Land

Welcome to Part C. This is where you go to find out what went wrong. Being sent to Part C is like explaining where the state of Delaware is. Folks who have been raised in the West do not know where Delaware is, so it makes it really difficult to explain.

Part C is the statistician's "never-ever land" because you can't apply statistics to a single transect or plot. So if you have only one transect or plot, you must do some more sampling work if you want to use statistics.

If you sampled just one plot, next time sample at least two or better yet three plots. If you sampled one frequency transect and tallied data, next time try to record data plot by plot or add more transects. Plotless techniques like the pace point or line intercept will require at least two and probably three or four transects to put statistical analysis to work.

Remember to always sample data within the mapping unit or area you are trying to inventory or monitor.

Pull another example or file that has multiple plots or transects and start again. If you don't have any, consider redesigning your monitoring program.

Part D - Detecting Change

Finally, we will use statistics to help us detect change over a period of time. Statistics can let us know our level of confidence and precision in stating that change has occurred. Unfortunately, there may be cases where change will be obvious, but the data collected do not statistically show it. Another thing to remember is that by the time statistics indicate change, rangeland may not respond to management. So "fine tune" (a nice verb for the TV generation, musicians, and auto mechanics) management as you go along.

Rule # 1 - In order to correctly state the statistical significance of a detected change, all studies must have been completed using the same ground rules and techniques each year. The number of plots or transects can vary, but it is never a good idea to do less than the baseline study.

Rule # 2 - Analysis of data must use the same analysis procedures each year; therefore, if you ever want to change analysis procedures, you must reanalyze all data using the same method.

Let's move on to more "productive range (ground)". You have already learned what the mean ($\bar{X}$) and variance (S^2) are. Now, we'll use the $\bar{X}$ and S^2 values from year 1 and year 2, as well as a value I'll have you look up in the t-table (Appendix 9) to detect change.

Find Figure D_1 and take it out so you can use it. As with Figures A_1 and B_1 there is a completed example on the right half of the page. The example has extra steps in the math computations. I just wanted to make sure you understand the order in which to multiply and add the values.

a. Fill in the blanks for n, $\bar{X}$, and S^2 for years 1 and 2. You should be able to copy them directly from forms A_1 or B_1. Notice that n, $\bar{X}$, and S^2 now have subscripts (a 1 indicates the earlier reading whereas a 2 indicates a subsequent or the latest reading) so that we can tell them apart.

b. Use the equation in line 1 to compute S^2, and enter it in the space provided. This is a new S^2 value. It is an average of S_1^2 and S_2^2, weighted by $(n_1-1)$and (n_2-1).

c. Use the equation in line 2 to compute S_d^2 (this confusing combination of letters and numbers is the variance of the difference between the earlier and later mean readings. Confused? . . . so am I!!. Put your value in the appropriate blank.

d. Use the equation in line 3 to compute the square root of S_d^2, and put it in the S_d blank.

e. Use the equation in line 4 to find df, the degrees of freedom you will use to look up the t-values in Appendix 9.

f. O.K. Using Appendix 9 and the df you have just calculated, fill in the blanks in line 5a with t-values. Now, multiply S_d (from line 3) by each of the t-values in line 5a and put the answers in line 5b.

g. Compute the difference between $\overline{X}_1$ and $\overline{X}_2$. Subtract the smaller from the larger so that diff (the difference) will be positive.

h. Finally, we're ready to see how confident we are that a change has occurred. Compare diff (from line 6) to the values in line 5b. If diff exceeds a value in line 5b, then you are 70% to 95% confident (depending on what the column heading is) that a change has occurred. Pick the largest value in line 5b which diff exceeds, and thus the corresponding confidence will be as high as possible.

O.K. We're through! Want a clue on how to set your objectives? You'll need $\overline{X}$, n, and S from a study you've already done.

a. First we'll have to look up a t-value in Appendix 9. Degrees of freedom will equal (2n)-2. Pick the t-value for the confidence level (e.g. 80% or 90%) you'll accept.

b. Compute the following: $t * S \div \sqrt{n}$

c. Add this value to $\overline{X}$ (your density, frequency, cover, etc. value). This is how large your subsequent $\overline{X}$ value must be to statistically exceed the first, original $\overline{X}$ at the confidence level you've selected. Remember this is only a target....an objective! If your variation for the first reading is different from the variation calculated for the second reading the figure may be a little off.

Year 1	Year 2
_____ n_1	_____ n_2
_____ $\overline{X}_1$	_____ $\overline{X}_2$
_____ S^2	_____ S^2

1. $[(n_1-1)S_1^2 + (n_2-1)S_2^2] \div (n_1+n_2-2) =$ _____ (S^2)

2. $(S^2 \div n_1) + (S^2 \div n_2) =$ _____ (S_d^2)

3. $\sqrt{S_d^2} =$ _____ (S_d)

4. $n_1 + n_2 - 2 =$ _____ (df)

5. t values from Appendix 9

Confidence Level	70%	80%	90%	95%
a. t value =	_____	_____	_____	_____
b. S_d x t =	_____	_____	_____	_____

6. $\overline{X}_1 - \overline{X}_2$ or $\overline{X}_2 - \overline{X}_1 =$ _____ (diff)
(diff must be zero or greater)

Conclusion:

Figure D_1

Year 1	Year 2
10 n_1	10 n_2
14 $\overline{X}_1$	16 $\overline{X}_2$
11.6 S^2	10.0 S^2

1. $[(n_1-1)S_1^2 + (n_2-1)S_2^2] \div (n_1+n_2-2)$ = _____ (S^2)
$[(9) \times 11.6 + (9) \times 10.0] \div (10 + 10 - 2) =$
$[104.4 + 90.0] \div 18 = \underline{10.80}\ (S^2)$

2. $(S^2 \div n_1) + (S^2 \div n_2) =$ _____ (S_d^2)
$10.80 \div 10 + 10.80 \div 10 =$
$1.08 + 1.08 = \underline{2.16}\ (S_d^2)$

3. $\sqrt{S_d^2} = \sqrt{2.16} = \underline{1.47}\ (S_d)$

4. $n_1 + n_2 - 2 = \underline{18}$ (df)

5.

Confidence Level	70%	80%	90%	95%
a. t value =	1.07	1.33	1.73	2.10
b. S_d x t =	1.57	1.96	2.54	3.09

6. $\overline{X}_1 - \overline{X}_2$ or $\overline{X}_2 - \overline{X}_1 = \underline{2.0}$ (diff)
(diff must be zero or greater)

Conclusion: $\overline{X}_2$ (our second reading) is greater than $\overline{X}_1$ (our first reading) at the 80% confidence level because 2.0 ($\overline{X}_2 - \overline{X}_1$) is greater than 1.96, but less than 2.54 (the value needed in order to be 90% confident).

Part E - Statistical Formulas

n = number of transects (frequency, line cover, point cover) or plots (plot cover, density, production)

X = species or sample values

$\overline{X}$ = mean or average formula: $\overline{X} = \frac{\Sigma X}{n}$

S^2 = variance formula: $S^2 = \frac{\Sigma(X - \overline{X})^2}{n-1}$

S = standard deviation formula: $S = \sqrt{S^2}$

CV = coefficient of variation formula: $CV = \frac{S}{\overline{X}}$

E = the precision or $\pm$ ___% of the mean. Note: in statistical calculations you always use the decimal equivalent i.e., $\pm$ 10% = $\pm$ .10

$t_{.90}$ = t value at the 90 % confidence level (or 10% chance to be wrong) level; (Note the decimal format) and where degrees of freedom = n-1. See Appendix 9 (t Table)

$t_{.95}$ = t value at 95% confidence (5% chance to be wrong)
$t_{.80}$ = t value at 80% confidence (20% chance to be wrong)

TO SOLVE FOR:

number (n) of required transects to reach a particular level of confidence and $\pm$% of the mean given t, CV, and E: $n = \frac{t^2 * CV^2}{E^2}$

$\pm$% of the mean (E) given t, CV, and n: $E = t * CV * \sqrt{1/n}$

coefficient of variation given n, E, and t: $CV = \frac{\sqrt{n} * E}{t}$

Note: t is shown without a particular level of confidence. The value of t must be obtained from the t table (Appendix 9).

ALLOTMENT: Plucked Goose
DATE: 9/31/84
OBSERVER: I. R. Quick
LOCATION: 1/4 mile W. of Dry Well #19
PARAMETER: Cover
TECHNIQUE: Daubenmire 6 class

% Cover (midpoint) By Plot

	1	2	3	4	5	6	7	8	9	10
Blackgrama	85.0	2.5	2.5	0.0	15.0	2.5	15.0	2.5	62.5	2.5
Western wheatgrass	62.5	2.5	15.0	15.0	37.5	2.5	2.5	15.0	87.5	2.5
Euphor(b)ia	2.5	15.0	0	0	2.5	15.0	0	0	2.5	15.0

ALLOTMENT: Cowtown
DATE: 10/07/84
OBSERVER: Down N. Out
LOCATION: 70 yards NW of truck
PARAMETER: Density
TECHNIQUE: 4 each 9.6 sq. ft. hoops

	Plot 1	Plot 2	Plot 3	Plot 4
Idaho Fescue	20	24	16	19
Phlox	5	7	3	6
Cheatgrass	70	60	65	69
Saguaro Cactus	1	0	0	0
Blue Spruce	1	0	0	0

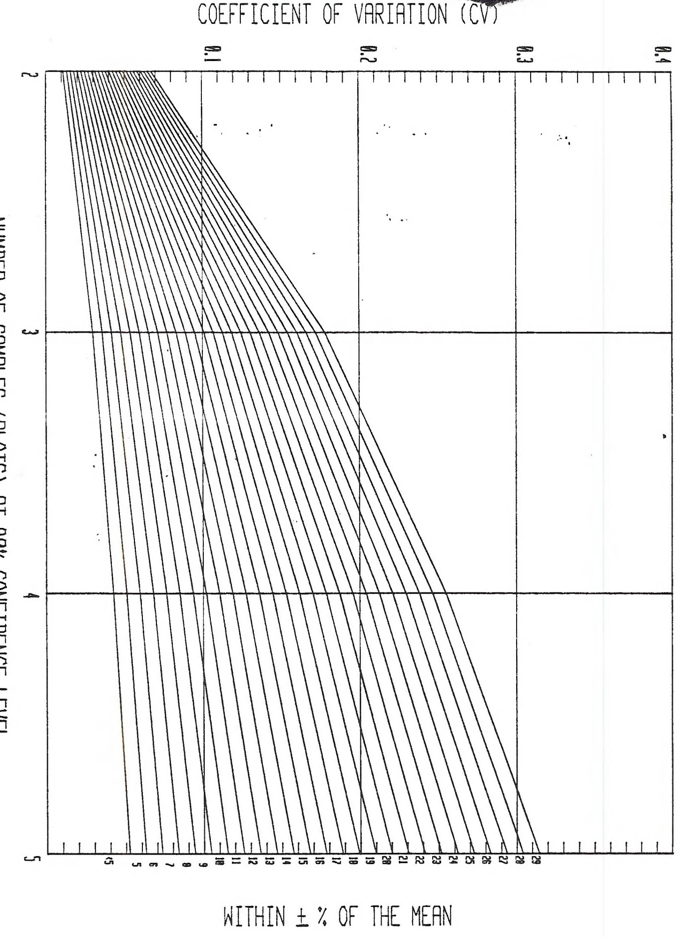
ALLOTMENT: Sheep Allotment 10
DATE: 11/31/89
OBSERVER: Where's d'Beef
LOCATION: 100 ft. SW of bedding ground
PARAMETER: Production.
TECHNIQUE: 10 9.6 sq. ft. plots (estimate of grams)

Plot #

	1	2	3	4	5	6	7	8	9	10
Globe Mallow	2	4	.0	0	1	10	0	4	0	1
Marsh Mallow	3	1	1	0	0	0	0	0	2	0
Barley	5	6	7	8	10	12	0	2	4	6
Single Leaf Pinyon	20	20	9	0	0	0	10	20	30	0
Squirreltail	9	10	8	12	5	3	10	0	0	1
Big Galleta	0	0	0	5	0	0	10	0	10	0
Cactus	5	0	0	0	0	5	0	0	0	0

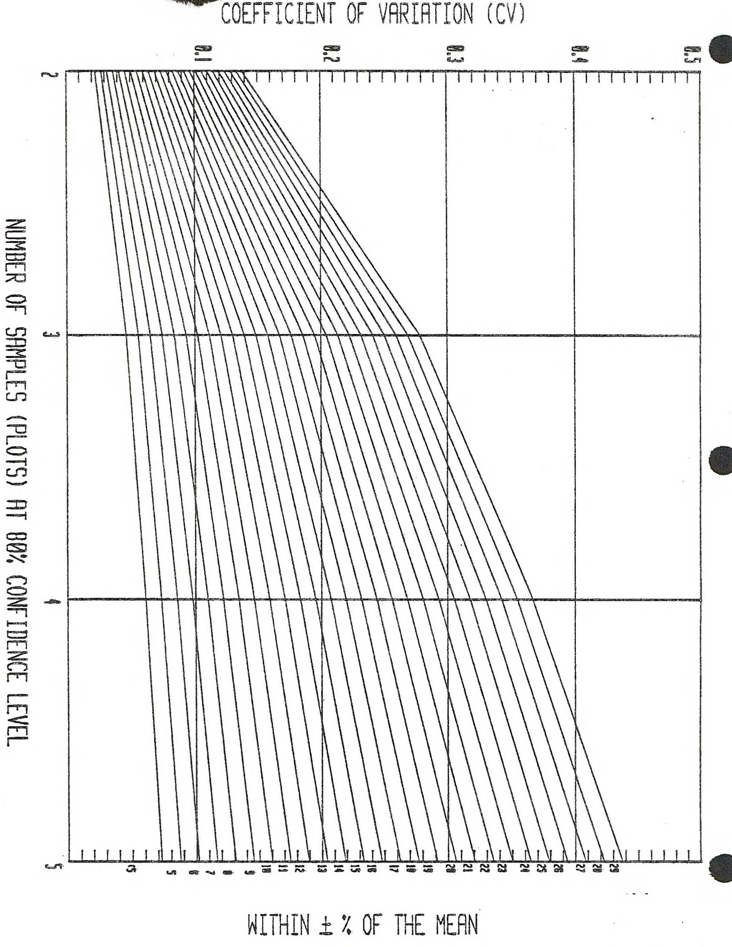
ALLOTMENT: Poverty Cattle Co.
DATE: Yesterday
OBSERVER: Ace Madrooga
LOCATION: 300 ft w of Go Broke Gulch sign
PARAMETER: Frequency
TECHNIQUE: 20 plots/subtransect - 5 subtransects

	Number Of Occurences Per Subtransect					Total %
	1	2	3	4	5	Frequency
Blue Bunch Wheatgrass	2	4	3	5	6	20%
Spike Muhly	9	10	11	9	10	49%
Filaree	16	17	16	19	20	88%
Western Red Cedar	1	1	0	2	0	4%
Nevada Bluegrass	10	12	13	9	13	57%
Burrograss	16	7	20	3	16	62%

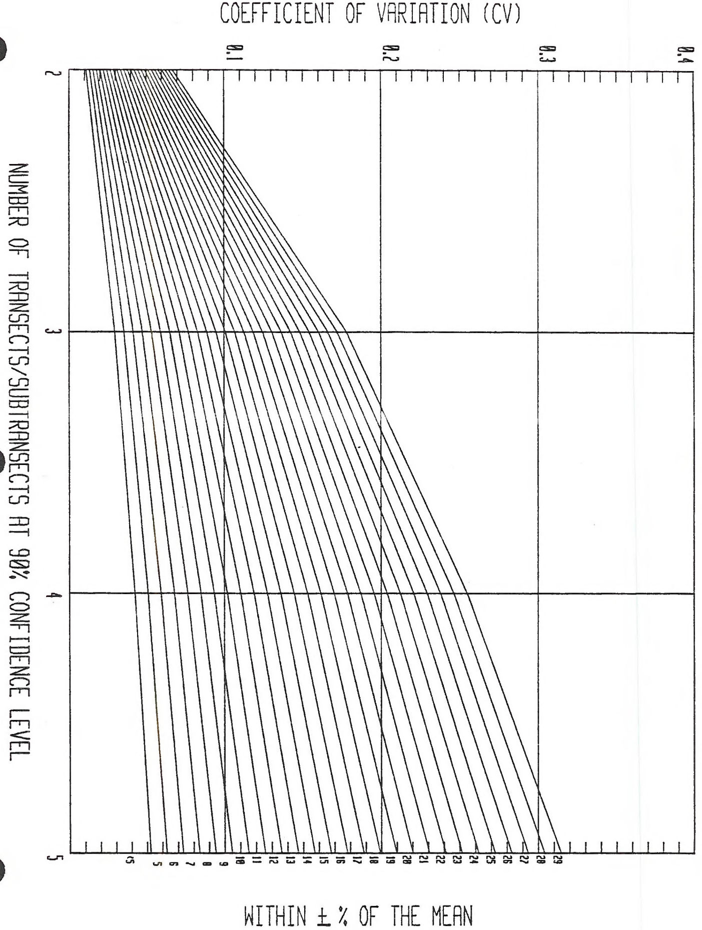


COEFFICIENT OF VARIATION (CV)
0.1
0.2
0.3
0.4
NUMBER OF SAMPLES (PLOTS) AT 90% CONFIDENCE LEVEL
2
3
4
5
5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29
WITHIN ± % OF THE MEAN

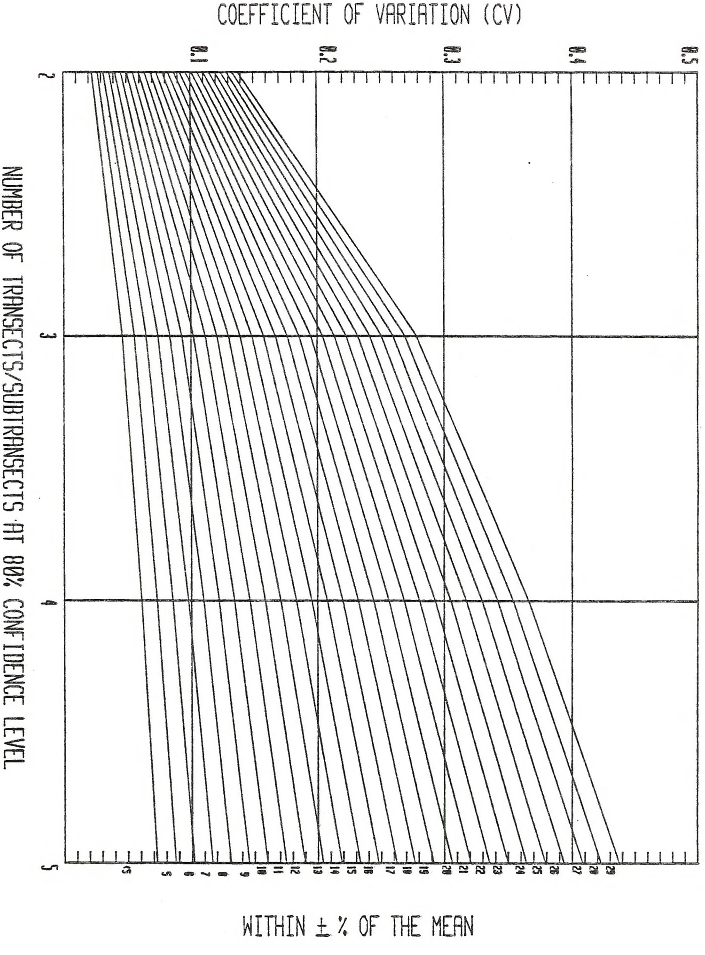
Appendix 5



COEFFICIENT OF VARIATION (CV)
0.1
0.2
0.3
0.4
0.5
NUMBER OF SAMPLES (PLOTS) AT 80% CONFIDENCE LEVEL
2
3
4
5
5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29
WITHIN ± % OF THE MEAN


COEFFICIENT OF VARIATION (CV)
0.1
0.2
0.3
0.4
NUMBER OF TRANSECTS/SUBTRANSECTS AT 90% CONFIDENCE LEVEL
2
3
4
5
<5 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29
WITHIN ± % OF THE MEAN

Appendix 7


COEFFICIENT OF VARIATION (CV)
0.1
0.2
0.3
0.4
0.5
NUMBER OF TRANSECTS/SUBTRANSECTS AT 80% CONFIDENCE LEVEL
2
3
4
5
<5
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
WITHIN ± % OF THE MEAN

Values of t

Level of Confidence

df	70%	80%	90%	95%
1	1.96	3.08	6.31	12.71
2	1.39	1.89	2.92	4.30
3	1.25	1.64	2.35	3.18
4	1.19	1.53	2.13	2.78
5	1.16	1.48	2.02	2.57
6	1.13	1.44	1.94	2.45
7	1.12	1.42	1.90	2.37
8	1.11	1.40	1.86	2.31
9	1.10	1.38	1.83	2.26
10	1.09	1.37	1.81	2.23
11	1.09	1.36	1.80	2.20
12	1.08	1.36	1.78	2.18
13	1.08	1.35	1.77	2.16
14	1.08	1.35	1.76	2.15
15	1.07	1.34	1.75	2.13
16	1.07	1.34	1.75	2.12
17	1.07	1.33	1.74	2.11
18	1.07	1.33	1.73	2.10
19	1.07	1.33	1.73	2.09
20	1.06	1.33	1.73	2.09
21	1.06	1.32	1.72	2.08
22	1.06	1.32	1.72	2.07
23	1.06	1.32	1.71	2.07
24	1.06	1.32	1.71	2.06
25	1.06	1.32	1.71	2.06
26	1.06	1.32	1.71	2.06
27	1.06	1.31	1.70	2.05
28	1.06	1.31	1.70	2.05
29	1.06	1.31	1.70	2.05
30	1.06	1.31	1.70	2.04
40	1.05	1.30	1.68	2.02
60	1.05	1.30	1.67	2.00
120	1.04	1.29	1.66	1.98
∞	1.04	1.28	1.65	1.96

Appendix 9

This certifies that:

knows as much about statistics as
anyone in this office and should
only be ~~insulted~~ consulted in
case of an emergency!

signed this day of ,19

ima Notta Numbercruncher
IMA NOTTA NUMBERCRUNCHER CHIEF FIGURER

Species ____________________ Date ____________ State ____________

Attribute __________________ Dist. ____________

Number of Plots, Frames, or Hoops ____________ (n) Allot. ____________

Study # ____________

	(X)	$-$	$(\bar{X})$	$=$	$(X-\bar{X})$	$\times$	$(X-\bar{X})$	$=$	$(X-\bar{X})^2$
Plot 1	____	-	____	=	____	x	____	=	____
Plot 2	____	-	____	=	____	x	____	=	____
Plot 3	____	-	____	=	____	x	____	=	____
Plot 4	____	-	____	=	____	x	____	=	____
Plot 5	____	-	____	=	____	x	____	=	____
Plot 6	____	-	____	=	____	x	____	=	____
Plot 7	____	-	____	=	____	x	____	=	____
Plot 8	____	-	____	=	____	x	____	=	____
Plot 9	____	-	____	=	____	x	____	=	____
Plot 10	____	-	____	=	____	x	____	=	____

TOTAL (X) ____ ÷ ____ (n) = ____ $(\bar{X})$ TOTAL A ____________

____ (TOTAL A) ÷ ____ (n-1) = ____ (S^2)

$\sqrt{\_\_\_\_ (S^2)}$ = ____ (S) ____ (S) ÷ ____ $(\bar{X})$ = ____ (CV)

INTERCEPT (n) and CV:

90% confidence ____ (% $\mp$ mean) 80% confidence ____ (% $\mp$ mean)

TO CALCULATE CONFIDENCE INTERVALS AROUND A MEAN OR DATA

@ 90% CONFIDENCE

(100%- ____ (% $\mp$ mean)) x ____ mean, density, cover (or production value) = lower limit ____

(100%+ ____ (% $\mp$ mean)) x ____ mean, density, cover (or production value) = upper limit ____

@ 80% CONFIDENCE

(100%- ____ (% $\mp$ mean)) x ____ mean, density, cover (or production value) = lower limit ____

(100%+ ____ (% $\mp$ mean)) x ____ mean, density, cover (or production value) = upper limit ____

Species ____________ Date ____________ State ____________
Attribute ____________ Dist. ____________
Number of transects (or subtransects) ______ (n) Allot. ____________
Study # ____________

	(X)	-	($\bar{X}$)	=	($X-\bar{X}$)	x	($X-\bar{X}$)	=	$(X-\bar{X})^2$
Transect 1	____	-	____	=	____	x	____	=	____
Transect 2	____	-	____	=	____	x	____	=	____
Transect 3	____	-	____	=	____	x	____	=	____
Transect 4	____	-	____	=	____	x	____	=	____
Transect 5	____	-	____	=	____	x	____	=	____
Transect 6	____	-	____	=	____	x	____	=	____
Transect 7	____	-	____	=	____	x	____	=	____
Transect 8	____	-	____	=	____	x	____	=	____
Transect 9	____	-	____	=	____	x	____	=	____
Transect 10	____	-	____	=	____	x	____	=	____

TOTAL (X) ____ ÷ ____ (n) = ____ ($\bar{X}$) TOTAL A ____________

____ (TOTAL A) ÷ ____ (n-1) = ____ (S^2)

$\sqrt{\_\_\_\_ (S^2)}$ = ____ (S)   ____ (S) ÷ ____ ($\bar{X}$) = ____ (CV)

INTERCEPT (n) and CV:

90% confidence ____ (% ∓ mean)   80% confidence ____ (% ∓ mean)

TO CALCULATE CONFIDENCE INTERVALS AROUND A MEAN OR DATA

@ 90% CONFIDENCE

(100%- ____ (%∓ mean)) x ____ (mean or frequency value) = lower limit ____

(100%+ ____ (%∓ mean)) x ____ (mean or frequency value) = upper limit ____

@ 80% CONFIDENCE

(100%- ____ (%∓ mean)) x ____ (mean or frequency value) = lower limit ____

(100%+ ____ (%∓ mean)) x ____ (mean or frequency value) = upper limit ____